

WHAT IS CLAIMED IS:

1. A method for clustering by address, comprising:  
receiving a search query that includes one or more keywords;  
obtaining one or more geographical identifiers;  
identifying an area of interest based, at least in part, on the one or more geographical identifiers;  
identifying documents that are associated with addresses located within the area of interest;  
determining ones of the identified documents that match the one or more keywords as relevant documents;  
grouping the relevant documents into clusters based, at least in part, on the addresses located within the area of interest; and  
presenting the clusters.
2. The method of claim 1, wherein the geographical identifiers are received as part of the search query.
3. The method of claim 1, wherein the geographical identifiers are inferred independent of the search query.
4. The method of claim 1, wherein the one or more keywords relate to a business or organization.

5. The method of claim 4, wherein the one or more geographical identifiers include location-specific information that approximately identifies a location of the business or organization.

6. The method of claim 1, wherein the one or more geographical identifiers include at least one of a partial address, a partial telephone number, an entire address, and an entire telephone number.

7. The method of claim 1, wherein the identifying an area of interest includes:  
determining a geographic location based, at least in part, on the one or more geographical identifiers,  
determining a geographic center of the geographic location, and  
identifying locations within a certain distance of the geographic center as the area of interest.

8. The method of claim 7, wherein the identifying locations includes:  
determining a radius, and  
identifying the area of interest as a circle centered on the geographic center with the determined radius.

9. The method of claim 8, wherein the radius is one of a predetermined radius and a radius set based on a specificity of the one or more geographical identifiers.

10. The method of claim 8, wherein the radius is a user-configurable radius.
11. The method of claim 8, wherein the radius is dynamically set based, at least in part, on the one or more keywords.
12. The method of claim 1, wherein the identifying documents includes:  
accessing a database that associates documents from a repository of crawled documents to addresses associated with the documents.
13. The method of claim 1, further comprising:  
scoring the relevant documents based on at least one of a distance factor and a relevancy factor.
14. The method of claim 13, wherein the distance factor for one of the relevant documents refers to a distance that an address associated with the one of the relevant documents is from a geographic center of the area of interest.
15. The method of claim 13, wherein the relevancy factor for one of the relevant documents refers to at least one of a number of the one or more keywords present in the one of the relevant documents and how prominently the one or more keywords appear in the one of the relevant documents.

16. The method of claim 1, wherein the grouping the relevant documents into clusters includes:

forming a separate one of the clusters for each of the addresses located within the area of interest.

17. The method of claim 1, wherein the grouping the relevant documents into clusters includes:

identifying a first one of the addresses associated with a first one of the relevant documents,

determining one or more second ones of the relevant documents that are also associated with the first address, and

grouping the first relevant document and the one or more second relevant documents into a cluster.

18. The method of claim 1, wherein the grouping the relevant documents into clusters includes:

placing each of the relevant documents into at least one cluster.

19. The method of claim 1, wherein the grouping the relevant documents into clusters includes:

placing at least one of the relevant documents into a plurality of the clusters.

20. The method of claim 1, wherein the presenting the clusters includes:

generating scores for the relevant documents within each of the clusters, and  
sorting the relevant documents within each of the clusters based, at least in part, on the  
scores.

21. The method of claim 1, wherein the presenting the clusters includes:  
ranking the clusters based on at least one of a distance factor and a relevancy factor, and  
sorting the clusters based, at least in part, on the ranking.

22. The method of claim 21, wherein the distance factor for one of the clusters refers  
to a distance that an address associated with the one cluster is from a geographic center of the  
area of interest.

23. The method of claim 22, wherein the relevancy factor for one of the clusters  
refers to at least one of a number of the one or more keywords present in at least one of the  
relevant documents in the one cluster and how prominently the one or more keywords appear in  
at least one of the relevant documents in the one cluster.

24. The method of claim 21, wherein the presenting the clusters further includes:  
weighting the distance factor and the relevancy factor differently based, at least in part,  
on the search query.

25. The method of claim 1, wherein the presenting the clusters includes:

forming a result output for each of the clusters, the result output including at least one of a title and a snippet for one of the relevant documents in the cluster and a title for another one or more of the relevant documents in the cluster.

26. The method of claim 1, wherein the presenting the clusters includes:

forming a result output for each of the clusters, the result output including a name of a business or organization and a title for one or more of the relevant documents in the cluster.

27. A system for forming search results, comprising:

means for receiving a search query;

means for identifying a geographical location;

means for determining a geographical center of the geographical location;

means for identifying locations within a certain distance of the geographical center as a geographical area of interest;

means for identifying documents that are associated with addresses located within the geographical area of interest; and

means for determining relevant ones of the identified documents, as relevant documents, based, at least in part, on the search query, the relevant documents forming the search results.

28. A system for forming search results, comprising:

a memory configured to store information that matches documents to addresses associated with the documents; and

a processor connected to the memory and configured to:

receive a search query,  
determine a geographical area of interest based, at least in part, on the search query,  
identify documents that are associated with addresses located within the geographical area of interest based, at least in part, on the information stored in the memory,  
group the identified documents into clusters based, at least in part, on the addresses located within the geographical area of interest, and  
provide the clusters as the search results.

29. A method for clustering by address, comprising:  
receiving a search query;  
identifying a geographical area of interest based, at least in part, on the search query;  
identifying documents that are associated with addresses located within the geographical area of interest;  
grouping the identified documents into clusters based, at least in part, on the addresses located within the geographical area of interest; and  
presenting the clusters.

30. A method for forming search results, comprising:  
receiving a search query that includes at least one portion of a telephone number;  
identifying a geographical area of interest based, at least in part, on the at least one portion of the telephone number;

identifying documents that are associated with addresses located within the geographical area of interest;

grouping the identified documents into clusters based, at least in part, on the addresses located within the geographical area of interest; and

presenting the clusters as the search results.

31. The method of claim 30, wherein the at least one portion of the telephone number includes at least one of an area code and a prefix associated with the telephone number.

32. A method for forming search results, comprising:

receiving a search query that includes one or more keywords and at least one portion of a telephone number;

identifying documents that are associated with telephone numbers that match the at least one portion of the telephone number;

determining ones of the identified documents that match the one or more keywords as relevant documents;

grouping the relevant documents into clusters based on the telephone numbers included in the relevant documents; and

presenting the clusters as the search results.

33. A method for forming search results, comprising:

receiving a search query that includes at least one portion of a telephone number;



identifying documents that are associated with telephone numbers that match the at least one portion of the telephone number;

grouping the identified documents into clusters based on the telephone numbers included in the identified documents; and

presenting the clusters as the search results.